Received: 19 August 2019 | Accepted: 4 December 2019

DOI: 10.1002/hyp.13678

RESEARCH ARTICLE

WILEY

# Event-based classification for global study of river flood generating processes

Lina Stein | Francesca Pianosi | Ross Woods

Department of Civil Engineering, University of Bristol, Bristol, UK

**Correspondence**
Lina Stein, Department of Civil Engineering, University of Bristol, Bristol, UK.
Email: lina.stein@bristol.ac.uk

**Funding information**
Engineering and Physical Sciences Research Council, Grant/Award Number: EP/L016214/1

## Abstract

Better understanding of which processes generate floods in a catchment can improve flood frequency analysis and potentially climate change impacts assessment. However, current flood classification methods are either not transferable across locations or do not provide event-based information. We therefore developed a location-independent, event-based flood classification methodology that is applicable in different climates and returns a classification of all flood events, including extreme ones. We use precipitation time series and very simply modelled soil moisture and snowmelt as inputs for a decision tree. A total of 113,635 events in 4155 catchments worldwide were classified into one of five hydro-climatological flood generating processes: short rain, long rain, excess rainfall, snowmelt and a combination of rain and snow. The new classification was tested for its robustness and evaluated with available information; these two tests are often lacking in current flood classification approaches. According to the evaluation, the classification is mostly successful and indicates excess rainfall as the most common dominant process. However, the dominant process is not very informative in most catchments, as there is a high at-site variability in flood generating processes. This is particularly relevant for the estimation of extreme floods which diverge from their usual flood generation pattern, especially in the United Kingdom, Northern France, Southeastern United States, and India.

**KEYWORDS**

classification, flood, flood frequency, flood generating process, flood type, global, large sample, mechanism

## 1 | INTRODUCTION

River flooding is a globally occurring natural hazard that takes many lives and causes extensive damage to property and infrastructure each year. Flood risk is predicted to increase in future years in several areas, particularly in Asia, Africa, and South America (Arnell & Gosling, 2016; Hirabayashi et al., 2013). Climate change, population growth, and urbanization all increase flood hazard and exposure (Hirabayashi et al., 2013). Large sample catchment studies already reveal historic trends in magnitude and frequency of floods over the past five to six decades in several areas around the world (Blöschl et al., 2019; Gudmundsson, Leonard, Do, Westra, & Seneviratne, 2019; Mallakpour & Villarini, 2015; Petrow & Merz, 2009). However, these trends in flood magnitude are not ubiquitous (Petrow & Merz, 2009)

**Abbreviations:** AWC, available water storage capacity; DFO, Dartmouth Flood Observatory; GLEAM, Global Land Evaporation Amsterdam Model; GSIM, Global Streamflow and Metadata Archive; MSWEP, Multi-Source Weighted-Ensemble Precipitation.

and cannot simply be connected to changes in precipitation (Sharma, Wasko, & Lettenmaier, 2018). Soil moisture or catchment wetness state often play an important role in flood generation (Berghuijs, Harrigan, Molnar, Slater, & Kirchner, 2019; Ivancic & Shaw, 2015; Sharma et al., 2018; Slater & Villarini, 2016). For example, Blöschl et al. (2017) showed that changes in the seasonal timing of extreme precipitation are not always the clearest explanatory factor for changes in the timing of floods, as in large areas of Europe flood occurrences are more influenced by timing of snowmelt or soil moisture maxima. This shows that rainfall alone is not the only driver of floods and river peak flow events can occur for multiple reasons. In addition to hydro-climatological processes (rainfall, snowmelt, rainfall excess on saturated ground, rain-on-snow), floods can be generated through blockages (e.g. ice jam, dam break) or tidal surges as well (Whitfield, 2012).

Information about flood generating processes can be used in a number of different ways. It can be used to explain detected trends in flood magnitude or timing (Gudmundsson et al., 2019; Mallakpour & Villarini, 2015; Petrow & Merz, 2009; Villarini & Slater, 2017) or to improve flood frequency analysis. In fact, the commonly used approach to flood frequency analysis assumes the flood sample to stem from a uniform distribution of flood events (England et al., 2018). However, it has been shown early on in several local studies that different flood processes generate different distributions and that using mixed distributions can improve flood frequency estimates (Elliott, Jarrett, & Ebling, 1982; Hirschboeck, 1987; Merz & Blöschl, 2008; Potter, 1958; Tarasova et al., 2019; Waylen & Woo, 1982). Furthermore, information about flood processes can contribute to improved flood modelling and forecasting (Viglione et al., 2010) and to flood risk management, as different flood types might generate different inundation behaviour (Sikorska, Viviroli, & Seibert, 2015).

For an in-depth overview of studies addressing the classification of flood events, see Tarasova et al. (2019). Here we focus on various studies that have identified flood generating processes based on hydro-climatological information within a catchment. Merz and Blöschl (2003) developed a widely used (e.g. Nied et al., 2014 ; Nied, Schröter, Lüdtke, Nguyen, & Merz, 2017 ; Sikorska et al., 2015) framework that uses combinations of catchment state (water stored in soil and snow) and climatic inputs to produce diagnostic maps at the regional and national scale for Austria. These maps show all process indicators for simultaneous events and allow the analyst to individually choose between one of five types: long rain floods, short rain floods, flash floods, rain-on-snow floods, and snowmelt floods (Merz & Blöschl, 2003). This approach allows an accurate description of flood events; however, one of the drawbacks is that the final flood type classification is not done automatically but left to the user. This is a subjective and time-consuming process that limits the number of events that can be considered. The other disadvantage is that the person interpreting the map has to be familiar with what constitutes, for example, a large rainfall amount for certain regions, since extreme rainfall amounts vary across Austria (Merz & Blöschl, 2003) and even more across larger regions (Boers et al., 2019).

Diezig and Weingartner (2007) advanced flood-type classification by introducing a decision tree to replace the user centred decisions. This concept was applied and extended also by Sikorska et al. (2015). The decision rules are still based on expert knowledge and the decision thresholds are based on literature values. Hydrograph information (Diezig & Weingartner, 2007) and process indicators based on weather (Diezig & Weingartner, 2007; Sikorska et al., 2015) and storage state (Sikorska et al., 2015) are inputs for the decision tree. Excluding discharge information can make this analysis available to catchments with limited data availability. Although the use of a decision tree makes this classification applicable to large samples and computationally efficient, the drawback of using regional thresholds based on literature values remains. It limits the applicability of their decision tree to regional studies, in this case in Switzerland, as set thresholds (e.g. event rainfall needs to be greater than 12 mm; Sikorska et al., 2015) are only valid in the intended region.

Larger scale studies that assess flood generating processes across different climates use different approaches. Berghuijs, Woods, Hutton, and Sivapalan (2016) determine dominant flood generating processes for the MOPEX catchments in the continental United States by comparing the mean occurrence date of the annual maximum peak flow to the occurrence date of the annual maximum of the hypothesized causal processes, which are: daily or multi-day precipitation events, precipitation excess or a combination of snowmelt and rain. Berghuijs et al. (2019) extend the seasonality statistics to Europe and additionally infer the relative distribution of different flood generating processes within one catchment by using circular statistics. Blöschl et al. (2017) analysed how flood timing in Europe changes based on observed changes in flood generating processes thus allowing a conclusion about which processes are most influential in certain regions. Blöschl et al. (2019) similarly found trends in flood magnitude to be closely related to changes both in precipitation and soil moisture for several areas in Europe. All these methods are based on the average timing of flood generating process versus average timing of flood event occurrence. This type of analysis determines dominant flood generating process; however, it cannot classify flood generating processes for individual events. While averaged results are still valuable, for example, for climate change impacts assessment (Blöschl et al., 2017), it still assumes the same flood processes within a catchment, which rarely is the case (Hirschboeck, 1987; Sikorska et al., 2015). The benefit of individual event process information is lost in such approaches. Any information pertaining particularly to extreme events, which might have a different generating process than the average annual maximum (Rogger et al., 2012; Smith, Cox, Baeck, Yang, & Bates, 2018), would not be available. Additionally, there are no studies that extend the analysis beyond the continental to the global scale, thus showing a lack in standardized flood process classification both for the dominant and the event scale analyses.

In this article, we present a widely transferable methodology to identify both the dominant flood generating process and the single-event generating process for catchments with different climates. The only streamflow information necessary for the analysis is the date of the annual maximum flow, thus reducing reliance on uncertain peak flow measurements. Other input data were kept to a minimum as well, so that a large number of stations could be included in the analysis,

and the method is potentially transferable to data-scarce regions. This is especially valuable, as any prior studies of flood generating processes have been focused on Europe or North America (Berghuijs et al., 2016; Berghuijs et al., 2019; Blöschl et al., 2017). Our study extends the knowledge of flood generating processes to other continents and climates as well. The new classification is tested for its robustness and evaluated with available information, two steps that are often lacking in current flood classification approaches (Tarasova et al., 2019).

## 2 | METHODOLOGY FOR A GLOBAL FLOOD CLASSIFICATION

This section describes how we infer flood generating processes at a global scale. Section 2.1 describes the flood event data source and how for each of the study catchments' daily rainfall, temperature, and evaporation data is calculated. These variables are then used as input to simple conceptual models producing daily snowmelt and soil moisture estimates (Section 2.2). Section 2.3 explains how flood process indicators are used in a conceptual decision tree to identify the flood generating process for a flood event. Section 2.5 then presents the methods applied to evaluate the robustness of the new classification.

### 2.1 | Data

The Global Streamflow Indices and Metadata Archive (GSIM) (Do, Gudmundsson, Leonard, & Westra, 2018a, 2018b; Gudmundsson, Do, Leonard, & Westra, 2018a, 2018b) provides streamflow station metadata, catchment delineation, catchment characteristics, and selected hydrological indices for more than 30,000 stations. While daily time series data are not made available, the date and magnitude of annual maximum flow are published. A value is supplied if at least 350 days of reliable daily flow data are available for that station (Gudmundsson et al., 2018b). For more information about the quality control both for the catchment delineation procedure and time series inhomogeneities refer to Gudmundsson et al. (2018b). Only stations which have a high quality in regard to both aspects, determined through the provided quality flags, are kept for the analysis. No Australian metadata were available to Do et al. (2018a) to quality check Australian catchments, we therefore used the catchment outlines delineated by Fowler, Peel, Western, Zhang, and Peterson (2016) to quality check the catchment delineation ourselves according to the criteria given by Do et al. (2018a). Out of 221 catchments in Fowler et al. (2016), 187 had a match in the GSIM database. In 62 catchments, the area between the catchment delineation by Do et al. (2018a) differed less than 10% from the catchments by Fowler et al. (2016) and were thus included in the analysis.

For the climate variables, global gridded climate data sets were chosen to make the analysis transferable across locations. Although global gridded data sets have location-specific uncertainties as well, the aim was to reduce uncertainty due to varying interpolation methods and to make the method potentially transferable into areas where precipitation gauge data are not available. A challenge in the selection of the data products was to find high-resolution data sets of global extent that cover several decades. For flood analysis, it is recommended to have at least 20 years of data available (Kjeldsen, 2015). Daily data are needed since a flood generating rainfall might just last for a few hours or days and would not be recognized in a coarser temporal resolution. This limits the choice of data sets available. For some data sets (e.g. temperature), a coarser resolution had to be accepted to be able to cover a longer time period.

The precipitation product used in the analysis is the Multi-Source Weighted-Ensemble Precipitation (MSWEP) Version 2.1 (Beck, Van Dijk, et al., 2017). It is a daily gridded precipitation product available at 0.1° x 0.1° resolution from 1979 to 2015 that merges satellite, reanalysis, precipitation gauge, and streamflow gauge data. Even without precipitation gauge correction, MSWEP is among the most accurate gridded precipitation products as evaluation with rainfall station data and hydrological modelling shows (Beck, Vergopolan, et al., 2017). MSWEP 2.1 additionally includes gauge correction. This means for some areas (snow-affected and/or 'complex topography'; Beck, Van Dijk, et al., 2017), streamflow station data are used to avoid underestimation of long-term precipitation estimates. Martens et al. (2017) use the MSWEP data as forcing data for the Global Land Evaporation Amsterdam Model (GLEAM) (Martens et al., 2017; Miralles, De Jeu, Gash, Holmes, & Dolman, 2011; Miralles, Holmes, et al., 2011). The current version GLEAM v3.2a is used as source for global actual evaporation information in this study. The model includes satellite, reanalysis, and merged products for various variables and provides daily evaporation data between 1980 and 2015 at 0.25° x 0.25° resolution. Although GLEAM has a tendency in some places to overestimate evapotranspiration (Khan, Liaqat, Baik, & Choi, 2018; Miralles et al., 2016), this fault is common among other evapotranspiration products as well. GLEAM is still more successful than comparable data products at closing the water balance (Miralles et al., 2016). For air temperature, the Berkeley Earth Surface Temperature daily gridded product was used (Rohde et al., 2013). It is available at 1° x 1° resolution from 1970 to 2013. It utilizes the Global Historical Climatology Network stations to interpolate a global grid of surface temperature. The large station network and interpolation algorithm by Rohde et al. (2013) produces a temperature product with lower uncertainties than comparable data (Menne et al., 2018). Although the daily gridded temperature product used here is not yet peer-reviewed, multiple sources assess it as comparable in accuracy to temperature products of similar extent and temporal resolution and it is used in multiple analyses (Levi Goss, 2013, e.g. Osborn, Jones, & Joshi, 2017; Wasko & Sharma, 2017). The different time extents of the above three data sources limit our analysis to the period 1980–2013.

### 2.2 | Snow and soil moisture accounting routine

For each climate variable, a catchment average daily value was produced using the GSIM delineated catchments. Each grid cell that is

covered by a catchment is assigned a weight based on the percentage of the catchment area that covers this cell. With these weights a weighted average value was calculated for each catchment. The catchment daily values were used as input into a simple coupled soil–snow routine to calculate snowmelt and soil saturation.

These simple models were adapted from the soil and snow routines used by Berghuijs et al. (2016). Snowmelt output is based on a degree-day snow model. Snow is accumulated if the temperature is below a critical temperature and melts if temperature is above. The daily time steps used did require some updates to the soil and snow routines applied by Berghuijs et al. (2016). One update is that rainfall input can only occur when the temperature is higher than the critical temperature thus separating rainfall and snowfall. The other is that the soil routine is only active if there is no snowpack, which is justified by the fact that soil moisture under a snowpack is not directly affected by precipitation and very little by evaporation. The bucket model gives an approximation of the spatially averaged moisture state of the catchment.

If $T(t) < T_{crit}$,

$$\begin{aligned} S_{snow}(t) &= S_{snow}(t-1) + P(t), \\ P_{melt}(t) &= 0, \\ P_{rain}(t) &= 0. \end{aligned} \tag{1}$$

If $T(t) > T_{crit}$,

$$\begin{aligned} P_{melt}(t) &= \min(f_{dd} * (T(t) - T_{crit}), S_{snow}(t-1)), \\ S_{snow}(t) &= S_{snow}(t-1) - P_{melt}(t), \end{aligned} \tag{2}$$

where at time $t$ $S_{snow}$ is snow storage (mm), $P$ precipitation input (mm), $P_{rain}$ is liquid precipitation, $T$ is air temperature (°C), $T_{crit}$ is the temperature threshold where rainfall turns to snow (°C), $f_{dd}$ is a melt factor (mm day$^{-1}$ K$^{-1}$), and $P_{melt}$ is snowmelt rate (mm). $f_{dd}$ is set to 2 mm day$^{-1}$ K$^{-1}$ (Berghuijs et al., 2019, 2016). For critical temperature, the data product by Jennings, Winchell, Livneh, and Molotch (2018) were used, which provide global gridded critical temperature for the Northern hemisphere. With no gridded data available for the Southern Hemisphere, critical temperature is set to 1°C, the mean critical temperature of the Northern hemisphere (Jennings et al., 2018). Unlike Berghuijs et al. (2016) the snow routine calculates only snowmelt and not rain-on-snow to separate these two processes. Although degree-day models are simplifications of snow processes, they work well at low resolution such as catchment averages (Hock, 2003) and are often used for flood classification (Tarasova et al., 2019). Snow storage was set to zero at the end of the annual average warmest month to include only annual snow and not accumulated snow over several years (Freudiger, Kohn, Stahl, & Weiler, 2014).

The information about snow storage ($S_{snow}$) is used as input for the soil routine, which assumes that soil filling, evaporation, and overflow only happens when there is no snow cover and when $T > T_{crit}$. Adapted from Berghuijs et al. (2016), soil storage and soil saturation are calculated as follows:

If $S_{snow}(t) = 0$,

$$\begin{aligned} S_u(t) &= \min(S_u(t-1) + P_{rain}(t), S_{u,max}) - Ea(t), \\ S_{sat} &= \frac{S_u(t)}{S_{u,max}}, \end{aligned} \tag{3}$$

$$P_{eff} = \max(S_u(t-1) + P_{rain}(t) - S_{u,max}, 0). \tag{4}$$

If $S_{snow}(t) > 0$,

$$\begin{aligned} S_u(t) &= S_u(t-1), \\ S_{sat} &= \frac{S_u(t)}{S_{u,max}}, \end{aligned} \tag{5}$$

$$P_{eff} = 0. \tag{6}$$

where $S_u$ is the soil storage (mm), $S_{u,max}$ is the maximum soil storage (mm), $S_{sat}$ is the soil saturation (%), $P_{eff}$ is the excess rainfall (mm), $P_{rain}$ is the liquid precipitation input (mm), and Ea is the actual evapotranspiration (mm). Maximum soil storage is set to the available water storage capacity (AWC) taken from the Harmonized World Soil Database averaged for each catchment (Nachtergaele, van Velthuizen, & Verelst, 2009). More detail including an evaluation of the combined soil–snow routine is given in Data S1, Supplementary A.

## 2.3 | Decision tree

A decision tree (Diezig & Weingartner, 2007; Sikorska et al., 2015) is used to decide which process generated each annual maximum flow event. While blockages (e.g. ice-jams, glacial outburst floods) can also cause extremely large floods, this article will only focus on hydro-climatological flood generating processes as they cause the overwhelming majority of floods (Whitfield, 2012). The structure of the tree is based on our domain knowledge. This means that the shape of the tree and the decision nodes are based on our understanding of flood generating processes, instead of being inferred from data through an automatic algorithm (Witten, Frank, & Hall, 2016). Specifically, several process indicators are calculated from the daily time series of precipitation, snowmelt, and soil moisture, which are then used by the tree to decide between five different flood generating processes: short rain flood, long rain flood, excess rainfall flood, rain/snowmelt flood, or snowmelt flood (Berghuijs et al., 2016; Merz & Blöschl, 2003; Sikorska et al., 2015). If none of these flood processes can be assigned, the flood event is described as 'other'. Flash floods and glacier melt floods are currently not included in the analysis.

The tree is used for classifying individual flood events whose dates of peak flow are known. The tree makes the classification decision based on information about potential causal factors (rain, snowmelt, soil moisture) that occur during a short time frame before the recorded flood peak. Since there is no daily discharge data available in GSIM for event separation as used by Sikorska et al. (2015), the time frame needs to be fixed. For alpine catchments, Froidevaux, Schwanbeck, Weingartner, Chevalier, and Martius (2015) found the time period relevant for flood generation is 3 days prior to a flood

event. Berghuijs et al. (2016) found for the MOPEX catchments in the United States that it varies between 3 and 10 days. We decided to set the threshold to 7 days in order to cover multi-day rain or snowmelt events even in large catchments. Without daily discharge time series, a more accurate delimitation is not possible. We evaluated the 7-day threshold by checking the relationship between catchment area and mean event response time (Figure S4). Based on the results, the time period of 7 days was considered applicable for both small and large catchments. With the flood relevant time period set to 7 days, the process indicators are each calculated for the 7-day period.

Each process indicator is used as a node in the decision tree. The decision tree with the process indicators is presented in Figure 1. Process indicators and associated thresholds are given in Table 1. A pseudocode description of the decision tree is given in Data S2, Supplementary B. The thresholds of the tree are inferred from the input time series itself by using either percentile thresholds (heavy rainfall, heavy snowmelt) or ratios (rain/snow). This makes the tree transferable across different locations. The only exception is the threshold for soil saturation, which is set to >90% (Sikorska et al., 2015), thus representing a near-saturated soil. Since this analysis focuses on annual maxima and not extreme floods, the 90th percentile is assumed to be a good indicator for finding large enough events to cause annual maximum flow.

Without using the Julian date of the flood occurrence, as was done by Sikorska et al. (2015), the other process indicators needed to be structured more strongly on hydrological reasoning. Hence, for our classification, any involvement of snow needs to be evaluated before all other processes, as it would be missed if soil moisture conditions or rainfall amount was evaluated first. Similarly, antecedent soil moisture conditions are evaluated before heavy rainfall events. In fact, Ivancic and Shaw (2015) demonstrated for the United States that extreme rainfall events are much more likely to cause an extreme streamflow event under wet antecedent conditions. Testing for heavy rainfall conditions

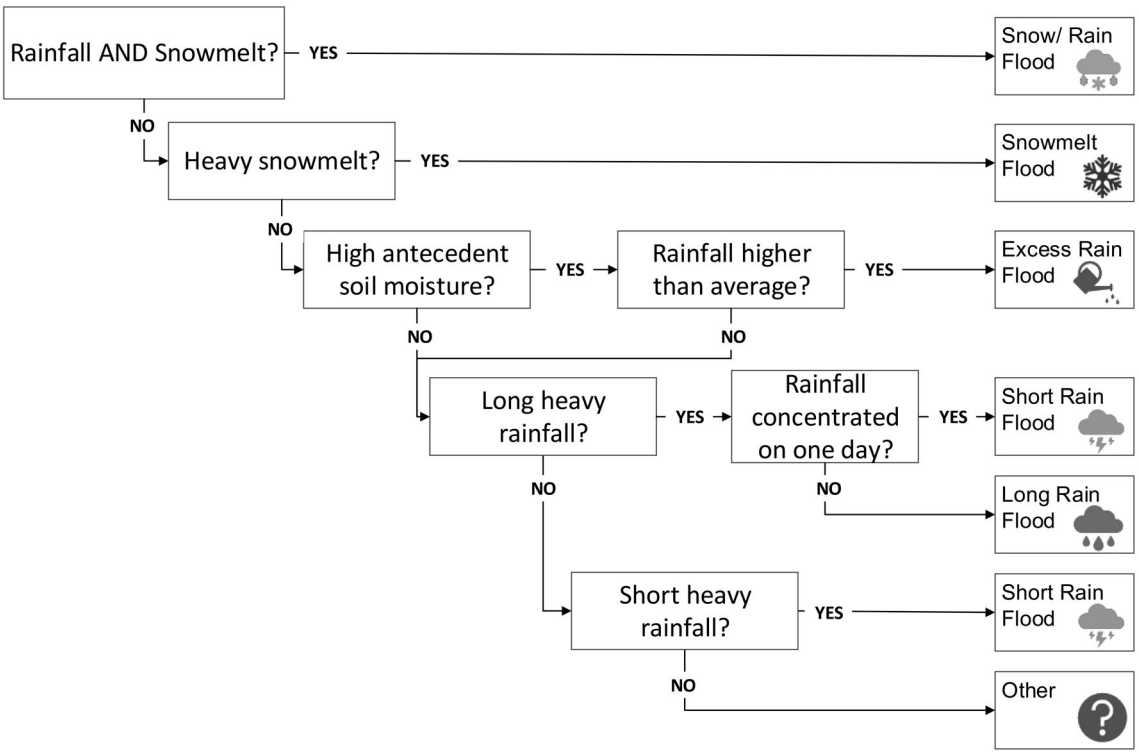


**FIGURE 1** Conceptual decision tree for location independent (global) flood classification. Thresholds are given in Table 1

**TABLE 1** Table of process indicators (rows) and their threshold values that lead to the different flood generating processes (columns)

| | | Snow/rain | Snowmelt | Excess rain | Short rain | Long rain |
|---|---|---|---|---|---|---|
| $P_{melt}(t_7)$ | Multi-day snowmelt | $>\frac{1}{3}P_{total}$ | $>p_{90}$ | | | |
| $P(t_7)$ | Multi-day rainfall | $>\frac{1}{3}P_{total}$ | | $>\bar{P}_7$ | $>p_{90}$ | $>p_{90}$ |
| $P(t)$ | Single-day rainfall | | | | $>p_{90}$ OR $>\frac{2}{3}P_7$ | |
| $S_{sat}$ | Soil moisture state | | | $>p_{90}$ | | |

*Note:* $p_{90}$ refers to the 90th percentile of the respective process indicator distribution. The snow/rain indicator is taken from Vormoor, Lawrence, Heistermann, and Bronstert (2015), and the soil moisture state threshold from Sikorska et al. (2015). $P_{melt}(t_7)$ is the snowmelt sum over 7 days before the flood event. $p_{90}$ for melt is calculated over all values $P_{melt}(t_7) > 1$. $P(t_7)$ is the rainfall sum over 7 days before the flood event. $P_{total} = P(t_7) + P_{melt}(t_7)$. $\bar{P}_7$ is the mean 7-day rainfall.

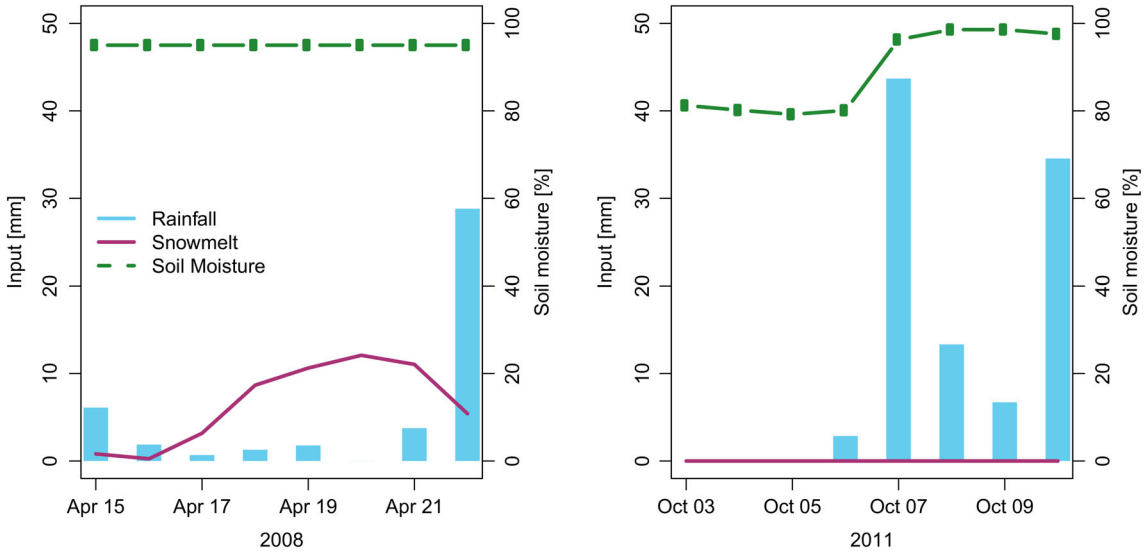


**FIGURE 2** Visual demonstration of input into the decision tree with rainfall, snowmelt and soil moisture distribution before a flood event (flood event occurred on the last day shown). The maximum peak flow for GSIM station number AT 0000032 (river Aschauer Ache in Austria) in 2008 (left) would be interpreted as a combination of rain and snow, the event in 2011 (right) as a long rainfall flood. GSIM, Global Streamflow and Metadata Archive

before excess rainfall conditions would therefore miss this important flood generating process. Nevertheless, the tree recognizes that wet conditions still require rainfall for flood generation (Berghuijs et al., 2019) by using an additional decision node, which tests if an above-average amount of rainfall fell in addition to wet antecedent conditions.

If none of these conditions are met, the category 'other' will be selected, which describes a flood event that is either misclassified or caused by something other than hydro-climatological processes (e.g. dam break, ice jam, groundwater flood, storm surge, etc.).

A visual demonstration of the inputs of the decision tree is given in Figure 2 for the 2008 and 2011 maxima in the Aschauer Ache catchment in Austria. In April 2008 (left), a melting event and a day of strong rainfall occurred right before the annual maximum flow, thus allowing the conclusion that the combination of snowmelt and rainfall generated the flood. The peak flow event in October 2011 (right) was instead preceded by several days of strong rainfall, falling on not yet wet soils. This would therefore be classified as a long rainfall flood.

## 2.4 | Dominant process and measure of process variability

A dominant flood generating process here is defined as the flood process occurring most often in the time series; however, in some catchments, this information is not as meaningful if flood generation varies evenly between two or more processes. The dominant process for each catchment can be identified if at least 20 years of events are classified in order to have a representative sample (Berghuijs et al., 2016; Kjeldsen, 2015).

The inter-annual variability of the flood generating process is calculated using a variability measure for categorical data taken from Allaj (2018):

$$v_k = 1 - \|f\|_k = 1 - \sqrt{f_0^2 + f_1^2 + \ldots + f_{k-1}^2} \quad (7)$$

where $v_k$ is the variability value, $k$ is the number of categories (flood generating processes), and $f$ is the relative frequencies for each category. When one process dominates, the value of $v_k$ is near 0; when several processes are equally dominant, the value of $v_k$ increases. If there are $k$ categories then the upper bound on $v_k$ is $1-1/\sqrt{k}$ (Allaj, 2018). The calculated variability is normalized to make the information easily accessible.

## 2.5 | Evaluation of the proposed global flood classification

### 2.5.1 | Evaluation by sensitivity analysis

A sensitivity analysis is used to determine the effect that changes in the inputs of the classification system may have on the output. It provides an insight into which input factors lead to a high variability of the output, thus evaluating robustness of the classification (Pianosi et al., 2016; Tarasova et al., 2019). In particular, the influence of the chosen model routine parameters and input data uncertainty (melt rate, AWC, critical temperature), as well as of the tree parameters and thresholds, is determined using a regional sensitivity analysis method (Spear & Hornberger, 1980; Young, Spear, & Hornberger, 1978). The tree structure is not evaluated with a sensitivity analysis as it is based on our hydrological understanding of flood processes. Changing the structure of the tree would produce outcomes not agreeing with our flood process definitions.

The flood classification is run with 1,000 parameter samples that are generated using a Latin hypercube sampling scheme over a uniform distribution. The range of parameters tested in the sensitivity

**TABLE 2** Table of initial parameter values of the model routine/decision nodes and the upper and lower parameter limits for the sensitivity analysis

| | Initial value | Lower limit | Upper limit |
|---|---|---|---|
| Melt rate (mm day $^{-1}$K$^{-1}$) | 2 | 1 | 8 |
| Uncertainty critical temperature (°C) | $T_{crit}$ | $T_{crit}$ − 1°C | $T_{crit}$ + 1°C |
| Uncertainty soil storage (%) | $S_{u,max}$ | $S_{u,max}$ − 50% | $S_{u,max}$ + 50% |
| Time period (days) | 7 | 3 | 14 |
| Rain/snowmelt overlap (–) | $\frac{1}{3}$ | $\frac{1}{10}$ | $\frac{1}{2}$ |
| Percentile heavy snow (–) | $p_{90}$ | $p_{80}$ | $p_{99}$ |
| Percentile heavy rain (–) | $p_{90}$ | $p_{80}$ | $p_{99}$ |
| Saturation threshold (%) | 90 | 90 | 99 |

*Note:* $T_{crit}$ values taken from Jennings et al. (2018), averaged by catchment. $S_{u,max}$ values taken from Nachtergaele et al. (2009), averaged by catchment.

analysis is displayed in Table 2. The parameter ranges are set to cover a plausible range of values. In particular, the catchment-averaged values of critical temperature and of maximum soil storage are varied by adding ±1°C and ±50% to the original value. Since the definition of excess rainfall floods requires saturated conditions, the saturation threshold is varied between 90 and 99%.

Following the rationale of the regional sensitivity analysis approach (Pianosi et al., 2016; Spear & Hornberger, 1980), the sensitivity to each parameter is determined by comparing the cumulative distributions of that parameter for each flood process to the uniform distribution of the entire parameter sample. The deviations from the uniform distribution are measured using the Kolmogorov–Smirnov statistic (*D* statistic). The lower these deviations, the less influence the parameter has on the classification outcome. Regional sensitivity analysis was applied using the SAFE toolbox (Pianosi, Sarrazin, & Wagener, 2015).

### 2.5.2 | Evaluation by comparison with available data

The accuracy of the event classification based on observed data is difficult to establish as flood types cannot be measured but depend on some other form of classification (Sikorska et al., 2015; Tarasova et al., 2019). Tarasova et al. (2019) state this as a disadvantage of current classification methods. For this reason, we evaluate our results with the only global scale data available for comparison, the Dartmouth Flood Observatory (DFO) Global Active Archive of Large Flood Events (Brakenridge, 2018) (http://floodobservatory.colorado.edu/Archives/index.html). A large flood as defined by the archive can be any event that received extensive media coverage, caused considerable damage, or resulted in fatalities (Brakenridge, 2018). The database provides event dates, an outline of the affected area, and cause of flood. The cause of flood included in the database is based on newspaper reports. We simplified and grouped flood cause into one of eight classes (Table S1) If any of the analysed events in our studies matched both in location and event date with the database entry, flood cause was compared to the output of our decision tree to evaluate the classification results. Since our study looks at annual maxima and not only at extreme events, a large percentage (96%) of the events we classified cannot be evaluated this way. Therefore, we additionally compare the results to various studies in the literature that describe flood generating processes for specific catchments and regions.

## 3 | RESULTS

### 3.1 | Dominant flood generating processes

Figure 3 shows the calculated dominant flood generating process, which is the process occurring most often in the time series. It also gives an overview of the distribution of the GSIM station locations used in this study. Although the station density varies, with the majority of stations in Northern America and Europe, many different regions and climates are covered so that a global analysis can be performed. In total, 4,155 catchments fulfilled the quality criteria specified by Do et al. (2018b) and Gudmundsson et al. (2018b) and had at least 20 years of data available for the analysis of the dominant process. In total, 113,635 events in total were classified. The prevalent dominant flood generating process for Brazil, Southeast Asia, India, most parts of Europe, the southeast of the United States, and New Zealand were classified as excess rainfall.

A region with distinctly different dominant process is South Africa and Namibia where most flood events were classified to occur during dry soil conditions, either due to short or long rainfall events. Similarly, the south of Switzerland has a mix of short rain and long rain floods, whereas further to the north, southeast of Germany is mainly classified as rain/snowmelt floods. Higher latitudes in Europe have more catchments with snowmelt and rain/snowmelt influenced floods, an exception being the coastline of Norway which is again dominated by excess rainfall. Dominance of excess rainfall floods decreases with decreasing storage capacity (Figure S5a).

The United States has a diverse landscape of dominant flood generating processes. While the southeast has widespread areas of floods generated by excess rainfall, there is a sharp longitudinal transition in

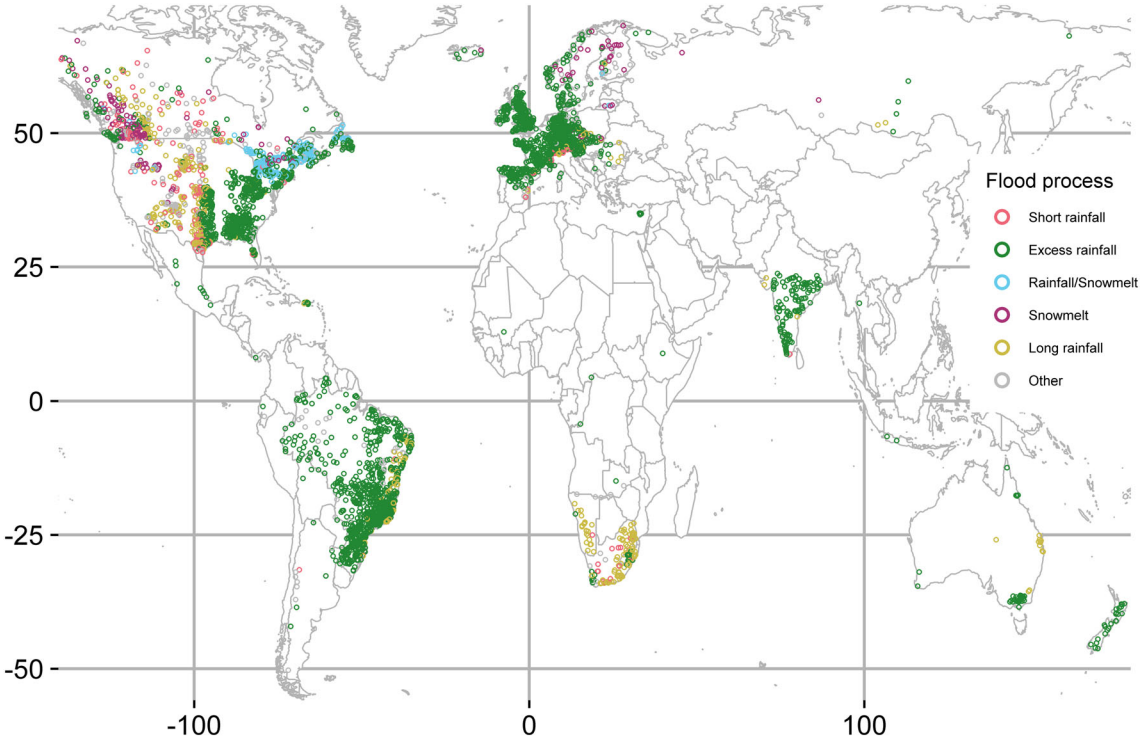


**FIGURE 3** Global map of dominant flood generating process for the 4,155 catchments of the study

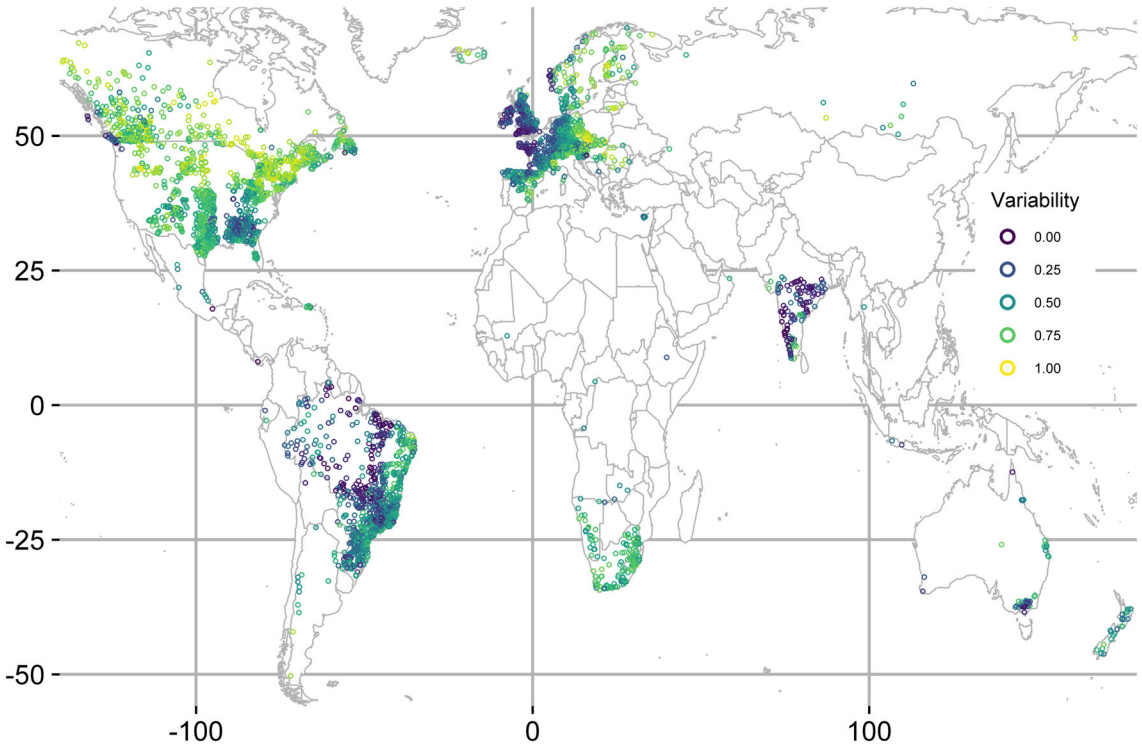


**FIGURE 4** Normalized year-to-year variability (Equation 7) of the flood generating process identified by our decision tree. A value of zero indicates the flood process is constant over the years, while high values (up to a maximum of one) indicate that different processes are identified in different years (Allaj, 2018)

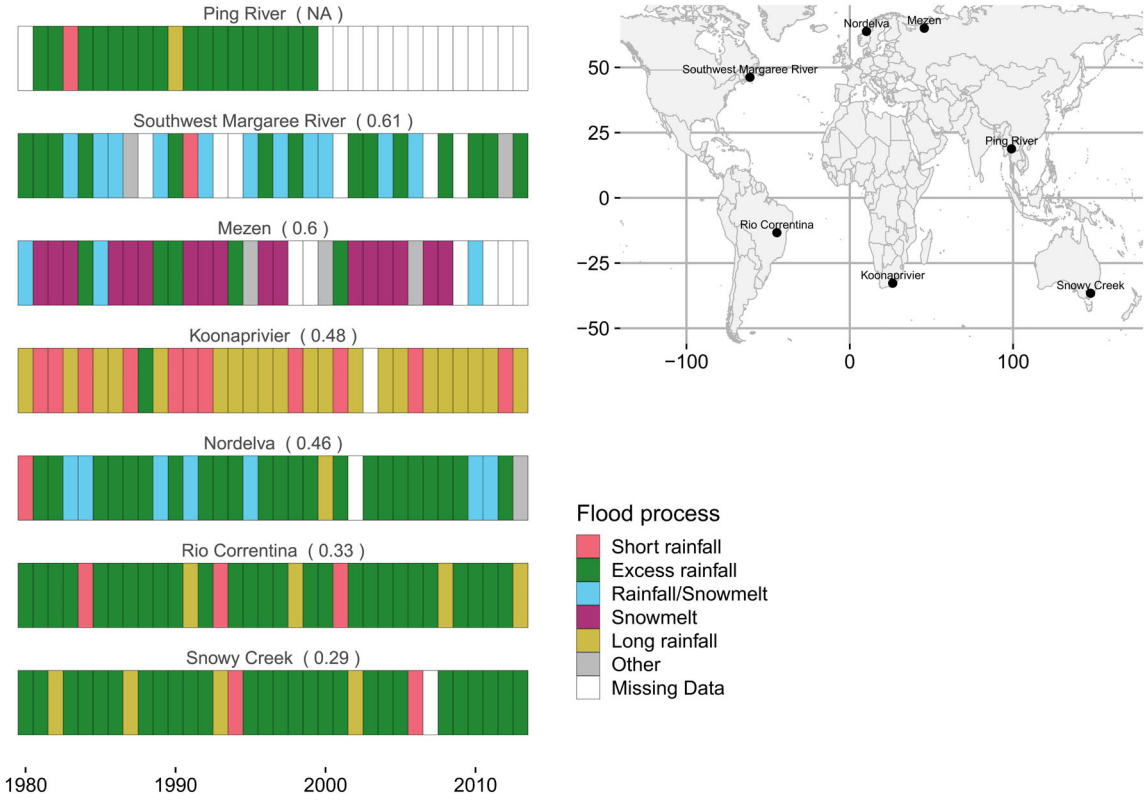


**FIGURE 5** Time series of annual flood generating process for seven example stations around the world illustrating the variability beyond dominant (i.e. most frequent) process. In brackets is the normalized inter-annual variability. The Ping river variability is missing as less than 20 years of data were available for calculation

the Southern United States towards short rain floods. The Northeastern United States is dominated by rainfall/snowmelt combinations. The Western United States do not have any clear patterns of a prevalent dominant flood generating process but are a mix between all four processes. Particularly in the south (New Mexico) and north (North Dakota, South Dakota), several catchments are dominantly classified as 'other'.

## 3.2 | Event-based flood generating processes

Figure 4 demonstrates the variability in flood generating processes in each catchment between different years. Within Europe, the United Kingdom and France have predominantly low variability in flood processes. India and Central/Northern Brazil are additional areas with low variability. The majority of catchments across different climates show a high variability (median normalized variability 0.59, mean 0.55), with some of the highest variability (normalized variability close to 1) reached in the Northern United States, Canada, and Central and Northern Europe. This demonstrates the need for more detailed information besides the dominant flood generating processes in these areas, as there might not be one dominant process but two or three processes combined.

For South Africa, variability is high (mean 0.65) as in most catchments short rain and long rain are almost equally often selected as the flood generating process. The exact distribution of flood generating processes for each catchment is shown in several detailed maps in Figure S5 provides an example of the annual flood generating processes time series for seven example catchments, for which detailed information in the literature is available.

The high variability in flood generating processes makes it quite likely that the flood generating process of the most extreme event in the time series is different to the dominant process. However, there are some areas, displayed in Figure 6, and some single catchments, where the process of the extreme flood is different, despite the usual flood generating process being very regular. The majority of the very stable catchments (variability < first quantile) where dominant flood process and extreme flood process are not the same, has excess rainfall as dominant process (796 out of 1,640 catchments compared to 212 for long rainfall, 209 for rain/snow, 179 for 'other', 152 for short rainfall, and 92 for snowmelt). The largest flood, in contrast, is caused by long rainfall (501 catchments), followed by short rainfall (246 catchments), or it remains unclear (class 'other', 87 catchments).

## 3.3 | Evaluation of the proposed global flood classification

### 3.3.1 | Evaluation by sensitivity analysis

Figure 7 shows the results of the sensitivity analysis. A high value of the Kolmogorov–Smirnov *D* statistic for a given parameter (horizontal

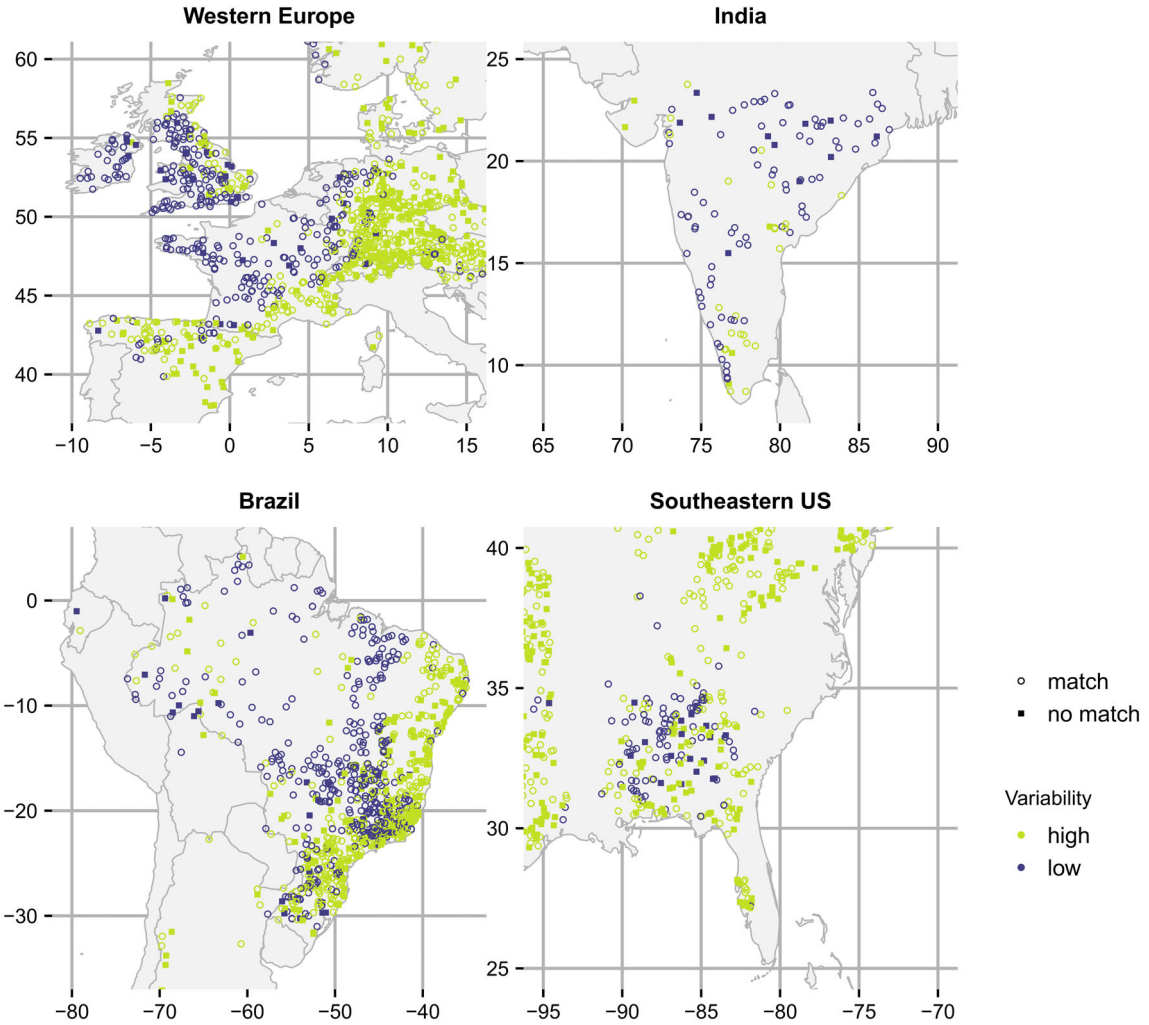


**FIGURE 6** Four regions where some catchments have an extreme flood generating process that does not match the dominant process (marked by a square), despite an overall low variability (marked in blue) in flood generating processes

axis) and flood process (colours) indicates strong sensitivity (Pianosi et al., 2016), that is, a change in that parameter has a strong effect on the number of times that process is classified. For the parameters of the soil–snow routine (melt rate, critical temperature, and maximum soil storage), the statistic is less than 0.1 for all processes. This indicates that a change in those values will not have a strong impact on the overall distribution of classified processes. The parameters for the simple model routine therefore have a low influence.

However, the classification of some flood processes is very sensitive to changes in the thresholds of the classification tree. The long rainfall and the 'other' outcomes are particularly sensitive to the time period threshold. The rain/snow outcome is strongly influenced by the rain/snowmelt overlap threshold. The snowmelt classification is most strongly influenced by choosing a different percentile threshold for heavy snowmelt. A change in the heavy rain threshold affects the classification of the outcomes long rainfall and 'other'. A change in the saturated conditions threshold affects events classified as excess rainfall. The higher the threshold, the fewer events of that process are classified (Figures 7 and S11). Figure S11 additionally informs us that a decrease of a threshold leads to an increase in the respective process. For the heavy rainfall threshold, this increase in classified long rainfall floods leads to a decrease in events classified as 'other'.

### 3.3.2 | Evaluation by comparison with available data

Figure 8 shows the comparison of the classified flood processes with the flood causes found by the Dartmouth Flood Observatory. About 4% of all classified events in our study have a match with an entry in the DFO database, both in timing and spatial overlap, as the database only takes by their definition large floods into account. Of those events, 5,101 events in total (84% of all matched events) are consistent with the global flood classification. The global flood classification detects well the rainfall component in flood generation (87% of all excess rainfall events classified as consistent, 93% of all short rainfall events classified as consistent, 90% of all long rainfall events classified as consistent). As newspaper reports do not take hydrological

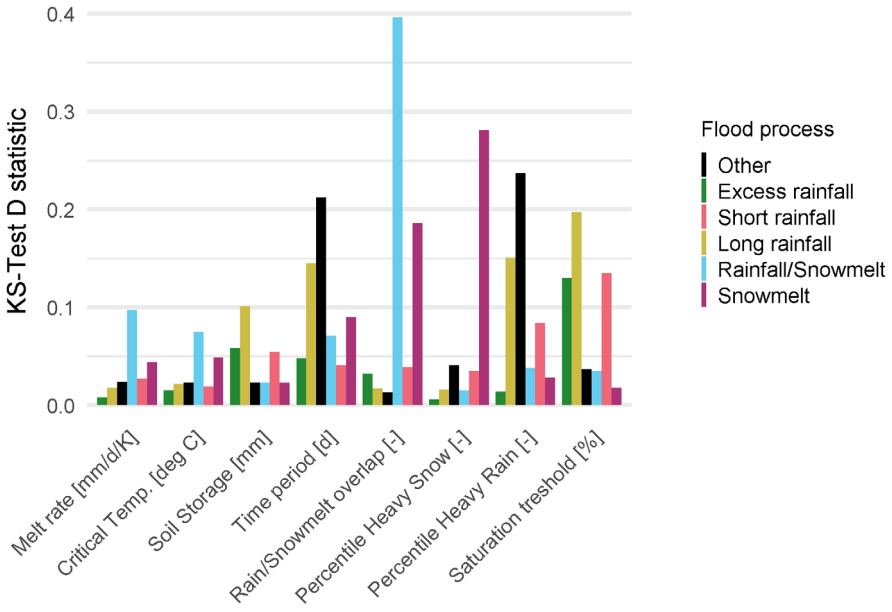


**FIGURE 7** *D* statistic for the Kolmogorov–Smirnov test for the regional sensitivity analysis. The higher the *D* statistic value, the more sensitive is the respective process to changes in the parameter/threshold. The theoretical range of *D* is between 0 and 1 (Pianosi & Wagener, 2015). The parameters of the model routine are melt rate, critical temperature, and soil storage. A more detailed figure with the cumulative distribution functions can be found in Figure S11

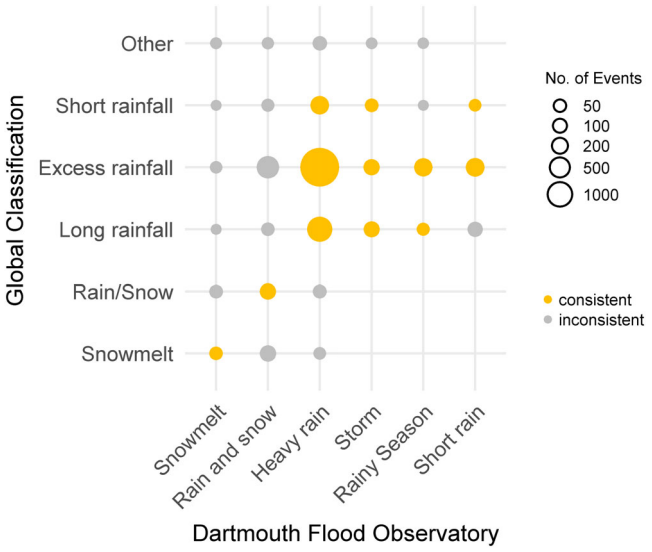


**FIGURE 8** Comparison of DFO flood causes with output of the global flood classification. Events are considered consistent if reported flood causes are plausible for news reports of the classified flood process (e.g. a DFO event where the cause is listed as rain is consistent with both soil moisture and rainfall floods). Yellow symbols indicate consistency between this study and DFO, and symbol size indicates number of flood events. An empty spot indicates zero flood events. DFO, Dartmouth Flood Observatory

processes such as soil moisture into account, a classification as excess rainfall can be considered consistent with the DFO classes 'Heavy rain', 'Storm', 'Rainy Season', and 'Short rain'. Large flood events with a snowmelt component seem to be less well classified (64% of all rain/snow events classified as consistent, 16% of all snowmelt events classified as consistent), as the very simple snowmelt routine in some cases will not accurately represent magnitude or timing of the snowmelt peak. This can cause rain/snow floods to be mistakenly classified only as rain floods.

## 4 | DISCUSSION

### 4.1 | Dominant flood generating processes

In some regions, the distribution of the dominant flood generating processes can be clearly linked to climate. For example, the distinct boundary between long/short rain and soil moisture floods in the Central–Southern United States (Figures 4 and S5b) is well mirrored in the aridity distribution in that area (e.g. Knoben et al., 2018). Arid and semi-arid regions rarely experience excess rainfall floods, and short rainfall and long rainfall are the prevalent generating processes there (Figure S5b). Examples for this are the more arid regions in South Africa, Namibia, and Australia. The only exception is seasonally arid catchments. We find excess rainfall to be a common flood process in seasonally arid catchments (Figure S5b).

In humid areas, especially in the humid tropics, we find excess rainfall to be the most common process. For the Ping catchment in Thailand (Figure 5), Lim and Boochabun (2012) describe that floods are not only due to tropical storms or monsoon rainfall but also require wet antecedent conditions. This demonstrates the benefit of focusing on flood processes instead of storm types (cyclone, monsoon, storm). Outside the tropics many studies describe soil moisture as a relevant factor in flood generation for several areas and river basins as well (Berghuijs et al., 2019; Institute of Hydrology (IoH), 1999; Lim & Boochabun, 2012). Our distribution of excess rainfall as dominant flood generating process matches previous studies in the United States (Berghuijs et al., 2016) and Europe (Berghuijs et al., 2019). However, Berghuijs et al. (2016) find no excess rainfall dominance in the Northeastern United States, whereas our classification

does. We also find that the dominant process is not very informative in that region due to the high inter-annual variability of flood processes. This might also explain why Berghuijs et al. (2016) could not find any dominant process for some of the catchments in that area. In Germany, there is a similar disagreement between Berghuijs et al. (2019) and our classification. Berghuijs et al. (2019) find most catchments in Western and Northern Germany to have no influence of snowmelt on flood generation. Our classification, however, finds that a small fraction of events in each catchment is caused by a combination of rain and snowmelt. Our findings are supported by Freudiger et al. (2014), who found several rain-on-snow events in the Rhine, Elbe, Weser, and Ems catchments, which cover most of Western and Northern Germany.

## 4.2 | Event-based flood generating processes

One key findings of this study is the widespread year-to-year variability in flood generating processes. Although some areas like Central Europe, India, and Central Brazil show low variability, the overwhelming majority of catchments are classified as regularly experiencing annual maxima generated by two or more different processes. This is very important for flood frequency analysis. Although it has long been known that frequency analysis, particularly of extreme floods, has higher accuracy if the flood distribution is split by flood type (Elliott et al., 1982; Hirschboeck, 1987; Potter, 1958; Waylen & Woo, 1982), such distinction is still not standard procedure. For example, the Bulletin 17-C for the United States does recommend separation of the flood frequency curve into different processes; however, it does not supply guidance on how to do this (England et al., 2018; Villarini & Slater, 2017), although there are recent approaches to rectify this (Barth, Villarini, & White, 2019). Especially in areas where the extreme flood process might deviate from the regular annual maxima, any flood estimation procedure might likely underestimate that extreme (Rogger et al., 2012; Smith et al., 2018). Areas where this may happen are, for example, the United Kingdom, the northwest of France, the southeast of the United States and Central India, as Figure 6 demonstrates. The detected generating processes for extreme events are dependent on the analysed time period (1980–2013). Although it is reasonable to expect that flood types change with time, questions about trends in flood type are beyond the scope of this paper. For some catchments the most extreme flood events might have happened outside that period and are therefore not included in this analysis. The classification shows that the shift in flood process mostly moves away from soil moisture towards long rainfall and short rainfall. This agrees with findings by Smith et al. (2018), who found unusually large floods in the United States are caused by a shift towards thunderstorms and tropical cyclones.

The benefit of an event-based classification is that it allows an event-based evaluation of the results. Events described as short rainfall in the Dartmouth Flood Observatory are classified by our global classification as either excess rainfall (possibly because wet antecedent conditions are not recorded in the Dartmouth Flood Observatory) or as long rainfall, which might be due to the limitation of using daily data. In fact, any overnight rainfall event would be registered as a 2-day event even if it only lasted few hours.

Most catchments in higher latitudes experience rain/snowmelt or snowmelt as flood generating processes. It rarely is the dominant process though with a few exceptions in mountainous areas or very high latitudes. According to the comparison with the Dartmouth Flood Observatory, not all snowmelt floods are classified correctly by our global flood classification. Rain and snow floods might get classified as snowmelt or vice versa. This is the case for the Southwest Margaree River (Figure 5) in Atlantic Canada, where a previous study by Collins et al. (2014) found that several catchments experience more rain floods than snowmelt floods. However, their methodology takes into account only the last 3 days prior to the flood event. Snowmelt and rain-on-snow floods can lead to a slow reaction of the catchment with snowmelt contributing to increased streamflow levels or a high wetness state (Merz & Blöschl, 2003). Therefore, 3 days might be too short to recognize the influence of snowmelt on flood generation. In contrast to Collins et al. (2014), Buttle et al. (2016) report that snow accumulation in the Maritime Provinces of Canada has a major impact on flood timing and magnitude, thus agreeing with the results reached by the global classification presented here. For comparison, in the Mezen catchment (Figure 5) our findings agree with the literature. In this catchment, Tockner, Uehlinger, and Robinson (2009) report snowmelt floods occur during the spring thaw, with additional occasional summer flash floods. In the Nordelva catchment (Figure 5), Vormoor et al. (2015) found 14% of flood events associated with rain-on-snow, with a dominance of 80% associated with rainfall. These proportions are similar to those found by our global flood classification. An evaluation of the outputs of the soil–snow routine (Figure S3) demonstrates that overall the storage simulations (snow accumulation and melt as well as water storing in the soil) work as expected. However, a possible reason why snowmelt processes could still not be perfectly represented in the classification is the potentially inaccurate timing of the snowmelt simulations. Catchment specific aspects and elevation variations cannot be taken into account with a fixed melt rate. A catchment with a wide range of elevations might have a mean temperature above the critical temperature, thus leading to melt conditions in the model routine, while in reality the snow pack is still present at higher elevations. This could lead to the model routine predicting a melt peak earlier or later than the actual peak, which means the snowmelt event is missed by the classification. One solution for this problem could be to define the model routine on a gridded basis instead of as lumped for each catchment. This would require downscaling gridded temperature to smaller grid cells which take topographic differences such as elevation into account.

## 4.3 | Classification tree method

A regional classification tree method (Diezig & Weingartner, 2007; Sikorska et al., 2015) was adapted and extended here to be transferable to several climates by using thresholds based on simulated time

series, instead of literature values. These climate independent thresholds make the tree applicable at global scale. As opposed to other flood classification methods (Merz & Blöschl, 2003; Nied et al., 2014; Sikorska et al., 2015), the classification tree does not require local knowledge of seasonality or weather patterns of the analysed catchments. The order of the decision nodes is based on hydrological process knowledge. An extension of the tree is possible, if further information is of interest. This could include glacier melt floods or a more detailed split of flood producing conditions.

The global extent of the application limits the temporal and spatial distribution of the data used as input to the classification. However, the methodology can easily be adapted to higher resolution data. This might be a higher temporal resolution of the climate input data or, if daily flow data is available, the classification of multiple flood events per year.

An important part of the decision tree is the last node 'other'. Including the class 'other' enables an evaluation of the original hypothesis of which flood generating processes should be included in the decision tree, and the identification of areas where that hypothesis fails. A region where most events are classified as 'other' is the Central United States. In North and South Dakota, this could be explained by ice jams, which are common during spring break up floods (McCabe & Crosby, 1959). Another location where that is the case is further to the south of the United States, towards New Mexico. A closer inspection of the rainfall distribution before flood events reveals that the flood generating rainfall is spread over 2 or 3 days. As the classification tree looks for either short (1 day) or long (7 days) extremes, these in-between storms might not reach the thresholds set. A more exact delimitation of the event time period, and thus a more exact delimitation of flood generating rainfall, might solve that problem. However, without daily streamflow data, a more accurate flood event delineation is difficult.

In their review, Tarasova et al. (2019) recommend to apply sensitivity analysis to test the robustness of the flood classification. With the exception of Sikorska et al. (2015), this is rarely done. Our sensitivity analysis revealed that the classification outcome is not very sensitive to the parameters of the soil–snow routine, but the choice of thresholds in the decision tree is. This result is consistent with what Sikorska et al. (2015) experienced with their crisp decision tree. A fuzzy decision tree might be considered a more robust approach; however, Sikorska et al. (2015) and Brunner et al. (2017) found that a fuzzy tree reached the same 'dominant' process per event as a crisp tree. An advantage of using a crisp tree is that it enabled the comparison with (crisp) flood cause data from the Dartmouth Flood Observatory.

This study extends river flood classification to a larger scale than has been done before. The focus of previous studies has been on continental (Europe), national (United States, Switzerland, Austria), or regional scale. One motivation to extend the analysis of flood generating processes to the global scale was to make the classification of flood generating processes comparable across more climates (Gupta et al., 2014; Linsley, 1982). Testing the classification across multiple climates and many catchments also highlights strengths and limitations of the methodology (Andréassian, Hall, Chahinian, & Schaake, 2006). For example, the events/catchments classified as 'other' are of interest. They reveal limitations of the classification method (as discussed above), but also point out locations where hydroclimatology is less influential for flood generation. This might be interesting in the context of ungauged catchment studies, because it can indicate either catchments that might not be suitable to be included in regionalization, or catchments having specific characteristics particularly interesting for regionalization (Boldetti, Riffard, Andréassian, & Oudin, 2010; Wagener & Wheater, 2006).

### 4.4 | Outlook

The widespread relevance of soil moisture in flood generation shown by this study can be applicable in climate change impact studies. Wasko and Sharma (2017) found changes in soil moisture with warming temperatures to be more relevant for streamflow response than extreme rainfall variation. The results presented here would support these findings, as extreme rainfall was not identified as the dominant process in most catchments. Nevertheless, it has to be taken into account that most catchments experience mixed processes and in some catchments where soil moisture is usually influential, a very extreme rainfall can still lead to extreme flooding. Therefore, if a catchment experiences occasional long rainfall/short rainfall floods, these might increase in frequency or magnitude, whereas excess rainfall floods might be less affected by changes in extreme rainfall. These different processes should be taken into account when flood risk changes are predicted (Slater & Villarini, 2017) and further studies regarding flood trends should focus on how floods of different processes might change differently. Additionally, the impact of climate change on flood generating processes will need to be examined in further studies. A first step in that direction could be a better understanding which catchment and climate characteristics are relevant in shaping the flood process mix.

## 5 | CONCLUSIONS

A new global methodology to analyse flood generating processes has been proposed and applied to 4,155 catchments of the GSIM database. Flood process indicators were queried in a decision tree to identify these processes for each annual flood peak flow event. The structure of the classification tree is dependent on the flood process definition and sensitive to changes in the threshold parameters. Nevertheless, the evaluation showed that most extreme flood events were classified consistent with reports from the Dartmouth Flood Observatory, with snowmelt influenced floods occasionally misclassified.

The analysis revealed that excess rainfall, that is, rainfall on wet soils, is a common flood generating process across several climates and continents. It also demonstrated the need for an event-based analysis, with a high variability of flood generating processes being the norm rather than the exception in most catchments. This should raise awareness of possible uncertainties in the common practice of using one distribution during flood frequency analysis to estimate

extreme floods. This is especially relevant since the most extreme and damaging floods might be generated by a process different from the dominant flood process (Rogger et al., 2012; Smith et al., 2018). The results found by the global flood classification are furthermore important for any future work analysing impact of system changes on flood events. Given the primary role of soil moisture in flood generation, the impact of predicted increases in extreme precipitation must be considered in the context of soil moisture, including future changes in soil moisture.

## ACKNOWLEDGEMENTS

This work was funded as part of the Water Informatics Science and Engineering Centre for Doctoral Training (WISE CDT) under a grant from the Engineering and Physical Sciences Research Council (EPSRC), grant number EP/L016214/1. The peak streamflow data used for the analysis are freely available from https://doi.pangaea.de/10.1594/PANGAEA.887477. We thank Hylke Beck for supplying the MSWEP data. We thank Conrad Wasko for supplying Australian catchment boundaries and Wouter Berghuijs, Larisa Tarasova, and Wouter Knoben for helpful discussions and feedback.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in 'global flood classification' at https://github.com/lshydro/global flood classification, Version 1. These data were derived from the following resources available in the public domain:

- Global Streamflow Indices and Metadata Archive (GSIM) https://doi.pangaea.de/10 .1594/PANGAEA.887477.
- Multi-Source Weighted-Ensemble Precipitation http://www.gloh2o.org/.
- Global Land Evaporation Amsterdam Model https://www.gleam.eu/.
- Berkeley Earth Surface Temperature http://berkeleyearth.org/.
- Harmonized World Soil Database v 1.2 http://www.fao.org/soils-portal/soil-survey/ soil-maps-and-databases/harmonized-world-soil-database-v12/en/.
- Gridded Northern Hemisphere 50% rain–snow $T_s$ threshold product https://doi.org/10.5061/dryad.c9h35.

## ORCID

*Lina Stein* https://orcid.org/0000-0002-9539-9549
*Francesca Pianosi* https://orcid.org/0000-0002-1516-2163
*Ross Woods* https://orcid.org/0000-0002-5732-5979

## REFERENCES

Allaj, E. (2018). Two simple measures of variability for categorical data. *Journal of Applied Statistics*, *45*(8), 1497–1516. https://doi.org/10.1080/02664763.2017.1380787

Andréassian, V., Hall, A., Chahinian, N., & Schaake, J. (2006). Why should hydrologists work on a large number of basin data sets? In V. Andréassian, A. Hall, N. Chahinian, & J. Schaake (Eds.), *Large sample basin experiments for hydrological model parameterization. Results of the model parameter experiment—MOPEX* (pp. 1–5). IAHS Press.

Arnell, N. W., & Gosling, S. N. (2016). The impacts of climate change on river flood risk at the global scale. *Climatic Change*, *134*(3), 387–401. https://doi.org/10.1007/s10584-014-1084-5

Barth, N. A., Villarini, G., & White, K. (2019). Accounting for mixed populations in flood frequency analysis: Bulletin 17C perspective. *Journal of Hydrologic Engineering*, *24*(3), 04019002. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001762

Beck, H., Van Dijk, A., Levizzani, V., Schellekens, J., Miralles, D. G., Martens, B., & de Roo, A. (2017). MSWEP: 3-Hourly 0.25 global gridded precipitation (1979–2015) by merging gauge, satellite, and reanalysis data. *Hydrology and Earth System Sciences*, *21*(1), 589–615. https://doi.org/10.5194/hess-21-589-2017

Beck, H., Vergopolan, N., Pan, M., Levizzani, V., Van Dijk, A., Weedon, G. P., … Wood, E. F. (2017). Global-scale evaluation of 22 precipitation datasets using gauge observations and hydrological modeling. *Hydrology and Earth System Sciences*, *21*(12), 6201–6217. https://doi.org/10.5194/hess-21-6201-2017

Berghuijs, W. R., Harrigan, S., Molnar, P., Slater, L. J., & Kirchner, J. W. (2019). The relative importance of different flood-generating mechanisms across Europe. *Water Resources Research*, *55*, 4582–4593. https://doi.org/10.1029/2019WR024841

Berghuijs, W. R., Woods, R. A., Hutton, C. J., & Sivapalan, M. (2016). Dominant flood generating mechanisms across the United States. *Geophysical Research Letters* https://doi.org/10.1002/2016GL068070, *43*, 4382–4390.

Blöschl, G., Hall, J., Parajka, J., Perdigão, R. A. P., Merz, B., Arheimer, B., … Chirico, G. B. (2017). Changing climate shifts of European floods. *Science*, *357*(6351), 588–590. https://doi.org/10.1126/science.aan2506

Blöschl, G., Hall, J., Viglione, A., Perdigão, R. A. P., Parajka, J., Merz, B., … Živković, N. (2019). Changing climate both increases and decreases European river floods. *Nature*, *573*(7772), 108–111. https://doi.org/10.1038/s41586-019-1495-6

Boers, N., Goswami, B., Rheinwalt, A., Bookhagen, B., Hoskins, B., & Kurths, J. (2019). Complex networks reveal global pattern of extreme-rainfall teleconnections. *Nature*, *566*(7744), 373–377. https://doi.org/10.1038/s41586-018-0872-x

Boldetti, G., Riffard, M., Andréassian, V., & Oudin, L. (2010). Data-set cleansing practices and hydrological regionalization: Is there any valuable information among outliers? *Hydrological Sciences Journal*, *55*(6), 941–951. https://doi.org/10.1080/02626667.2010.505171

Brakenridge, G. R. (2018). *Global active archive of large flood events*. Retrieved from University of Colorado website: http://floodobservatory.colorado.edu/Archives/index.html

Brunner, M. I., Viviroli, D., Sikorska, A. E., Vannier, O., Favre, A.-C., & Seibert, J. (2017). Flood type specific construction of synthetic design hydrographs. *Water Resources Research*, *53*(2), 1390–1406. https://doi.org/10.1002/2016WR019535

Buttle, J. M., Allen, D. M., Caissie, D., Davison, B., Hayashi, M., Peters, D. L., … Whitfield, P. H. (2016). Flood processes in Canada: Regional and special aspects. *Canadian Water Resources Journal*, *41*(1–2), 7–30. https://doi.org/10.1080/07011784.2015.1131629

Collins, M. J., Kirk, J. P., Pettit, J., DeGaetano, A. T., McCown, M. S., Peterson, T. C., … Zhang, X. (2014). Annual floods in New England (USA) and Atlantic Canada: Synoptic climatology and generating mechanisms. *Physical Geography*, *35*(3), 195–219. https://doi.org/10.1080/02723646.2014.888510

Diezig, R., & Weingartner, R. (2007). Hochwasserprozesstypen in der Schweiz. *Wasser Und Abfall*, *4*(1), 18–26. Retrieved from https://boris.unibe.ch/id/eprint/25512

Do, H. X., Gudmundsson, L., Leonard, M., & Westra, S. (2018a). *The global streamflow indices and metadata archive—Part 1: Station catalog and catchment boundary*. https://doi.org/10.1594/PANGAEA.887477

Do, H. X., Gudmundsson, L., Leonard, M., & Westra, S. (2018b). The global streamflow indices and metadata archive (GSIM)—Part 1: The production of a daily streamflow archive and metadata. *Earth System*

*Science Data*, *10*(2), 765–785. https://doi.org/10.5194/essd-10-765-2018

Elliott, J., Jarrett, R., & Ebling, J. (1982). *Annual snowmelt and rainfall peak-flow data on selected foothills region streams, South Platte River, Arkansas River, and Colorado River basins*. Lakewood, CO.

England, J. F. Jr., Cohn, T. A., Faber, B. A., Stedinger, J. R., Thomas, W. O. Jr., Veilleux, A. G., … Mason, R. R. Jr. (2018). Guidelines for determining flood flow frequency—Bulletin 17C. Techniques and Methods 4-B5. In *Hydrologic analysis and interpretation*. Reston, VA: U.S. Geological Survey. https://doi.org/10.3133/tm4b5

Fowler, K., Peel, M. C., Western, A. W., Zhang, L., & Peterson, T. J. (2016). Simulating runoff under changing climatic conditions: Revisiting an apparent deficiency of conceptual rainfall-runoff models. *Water Resources Research*, *52*(3), 1820–1846. https://doi.org/10.1002/2015WR018068

Freudiger, D., Kohn, I., Stahl, K., & Weiler, M. (2014). Large-scale analysis of changing frequencies of rain-on-snow events with flood-generation potential. *Hydrology and Earth System Sciences*, *18*(7), 2695–2709. https://doi.org/10.5194/hess-18-2695-2014

Froidevaux, P., Schwanbeck, J., Weingartner, R., Chevalier, C., & Martius, O. (2015). Flood triggering in Switzerland: The role of daily to monthly preceding precipitation. *Hydrology and Earth System Sciences*, *19*(9), 3903–3924. https://doi.org/10.5194/hess-19-3903-2015

Gudmundsson, L., Do, H. X., Leonard, M., & Westra, S. (2018a). *The global streamflow indices and metadata archive (GSIM)—Part 2: Time series indices and homogeneity assessment*. https://doi.org/10.1594/PANGAEA.887470

Gudmundsson, L., Do, H. X., Leonard, M., & Westra, S. (2018b). The Global Streamflow Indices and Metadata Archive (GSIM)—Part 2: Quality control, time-series indices and homogeneity assessment. *Earth System Science Data*, *10*(2), 787–804. https://doi.org/10.5194/essd-10-787-2018

Gudmundsson, L., Leonard, M., Do, H. X., Westra, S., & Seneviratne, S. I. (2019). Observed trends in global indicators of mean and extreme streamflow. *Geophysical Research Letters*, *46*(2), 756–766. https://doi.org/10.1029/2018GL079725

Gupta, H. V., Perrin, C., Blöschl, G., Montanari, A., Kumar, R., Clark, M. P., & Andréassian, V. (2014). Large-sample hydrology: A need to balance depth with breadth. *Hydrology and Earth System Sciences*, *18*(2), 463–477. https://doi.org/10.5194/hess-18-463-2014

Harris, I., Jones, P., Osborn, T., & Lister, D. (2014). Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 Dataset. *International Journal of Climatology*, *34*(3), 623–642.

Hirabayashi, Y., Mahendran, R., Koirala, S., Konoshima, L., Yamazaki, D., Watanabe, S., … Kanae, S. (2013). Global flood risk under climate change. *Nature Climate Change*, *3*(9), 816–821. https://doi.org/10.1038/nclimate1911

Hirschboeck, K. K. (1987). Hydroclimatically-defined mixed distributions in partial duration flood series. In *Hydrologic frequency modeling* (pp. 199–212). https://doi.org/10.1007/978-94-009-3953-0_13

Hock, R. (2003). Temperature index melt modelling in mountain areas. *Journal of Hydrology*, *282*(1–4), 104–115. https://doi.org/10.1016/S0022-1694(03)00257-9

Institute of Hydrology (IoH) (1999). *Flood estimation handbook*. Wallingford, UK.

Ivancic, T. J., & Shaw, S. B. (2015). Examining why trends in very heavy precipitation should not be mistaken for trends in very high river discharge. *Climatic Change*, *133*(4), 681–693. https://doi.org/10.1007/s10584-015-1476-1

Jennings, K. S., Winchell, T. S., Livneh, B., & Molotch, N. P. (2018). Spatial variation of the rain–snow temperature threshold across the Northern Hemisphere. *Nature Communications*, *9*(1), 1148. https://doi.org/10.1038/s41467-018-03629-7

Khan, M. S., Liaqat, U. W., Baik, J., & Choi, M. (2018). Stand-alone uncertainty characterization of GLEAM, GLDAS and MOD16 evapotranspiration products using an extended triple collocation approach. *Agricultural and Forest Meteorology*, *252*, 256–268. https://doi.org/10.1016/J.AGRFORMET.2018.01.022

Kjeldsen, T. R. (2015). How reliable are design flood estimates in the UK? *Journal of Flood Risk Management*, *8*(3), 237–246. https://doi.org/10.1111/jfr3.12090

Knoben, W. J. M., Woods, R. A., & Freer, J. E. (2018). A quantitative hydrological climate classification evaluated with independent streamflow data. *Water Resources Research*, *54*(7), 5088–5109. https://doi.org/10.1029/2018WR022913

Levi Goss, B. (2013). Earth's land surface temperature trends: A new approach confirms previous results. *Physics Today*, *66*(4), 17–19. https://doi.org/10.1063/PT.3.1936

Lim, H. S., & Boochabun, K. (2012). Flood generation during the SW monsoon season in northern Thailand. *Geological Society, London, Special Publications*, *361*(1), 7–20. https://doi.org/10.1144/SP361.3

Linsley, R. K. (1982). Rainfall–runoff models-an overview. In V. P. Singh (Ed.), *Proceedings of the international symposium on rainfall-runoff modeling* (pp. 3–22). Littleton, CO: Water Resources Publications.

Mallakpour, I., & Villarini, G. (2015). The changing nature of flooding across the Central United States. *Nature Climate Change*, *5*(3), 250–254. https://doi.org/10.1038/nclimate2516

Martens, B., Gonzalez Miralles, D., Lievens, H., Van Der Schalie, R., De Jeu, R. A., Fernández-Prieto, D., … Verhoest, N. (2017). GLEAM v3: Satellite-based land evaporation and root-zone soil moisture. *Geoscientific Model Development*, *10*(5), 1903–1925. Retrieved from https://biblio.ugent.be/publication/8520679

McCabe, J. A., & Crosby, O. A. (1959). *Floods in North and South Dakota, frequency and magnitude*.

Menne, M. J., Williams, C. N., Gleason, B. E., Rennie, J. J., Lawrimore, J. H., Menne, M. J., … Lawrimore, J. H. (2018). The global historical climatology network monthly temperature dataset, version 4. *Journal of Climate*, *31*(24), 9835–9854. https://doi.org/10.1175/JCLI-D-18-0094.1

Merz, R., & Blöschl, G. (2003). A process typology of regional floods. *Water Resources Research*, *39*(12). https://doi.org/10.1029/2002WR001952

Merz, R., & Blöschl, G. (2008). Flood frequency hydrology: 2. Combining Data Evidence. *Water Resources Research*, *44*(8). https://doi.org/10.1029/2007WR006745

Miralles, D. G., De Jeu, R. A. M., Gash, J. H., Holmes, T. R. H., & Dolman, A. J. (2011). Magnitude and variability of land evaporation and its components at the global scale. *Hydrology and Earth System Sciences*, *15*(3), 967–981. https://doi.org/10.5194/hess-15-967-2011

Miralles, D. G., Holmes, T. R. H., De Jeu, R. A. M., Gash, J. H., Meesters, A. G. C. A., & Dolman, A. J. (2011). Global land-surface evaporation estimated from satellite-based observations. *Hydrology and Earth System Sciences*, *15*(2), 453–469. https://doi.org/10.5194/hess-15-453-2011

Miralles, D. G., Jiménez, C., Jung, M., Michel, D., Ershadi, A., McCabe, M. F., … Fernández-Prieto, D. (2016). The WACMOS-ET project—Part 2: Evaluation of global terrestrial evaporation data sets. *Hydrology and Earth System Sciences*, *20*(2), 823–842. https://doi.org/10.5194/hess-20-823-2016

Nachtergaele, F., van Velthuizen, H., & Verelst, L. (2009). *Harmonized world soil database, Version 1.1*. Rome: FAO/IIASA/ISRIC/ISSCAS/JRC.

Nied, M., Pardowitz, T., Nissen, K., Ulbrich, U., Hundecha, Y., & Merz, B. (2014). On the relationship between hydro-meteorological patterns and flood types. *Journal of Hydrology*, *519*, 3249–3262. https://doi.org/10.1016/j.jhydrol.2014.09.089

Nied, M., Schröter, K., Lüdtke, S., Nguyen, V. D., & Merz, B. (2017). What are the hydro-meteorological controls on flood characteristics? *Journal of Hydrology*, *545*, 310–326. https://doi.org/10.1016/J.JHYDROL.2016.12.003

Osborn, T. J., Jones, P. D., & Joshi, M. (2017). Recent United Kingdom and global temperature variations. *Weather*, *72*(11), 323–329. https://doi.org/10.1002/wea.3174

Petrow, T., & Merz, B. (2009). Trends in flood magnitude, frequency and seasonality in Germany in the period 1951–2002. *Journal of Hydrology*, *371*(1–4), 129–141. https://doi.org/10.1016/J.JHYDROL.2009.03.024

Pianosi, F., Beven, K. J., Freer, J., Hall, J. W., Rougier, J., Stephenson, D. B., & Wagener, T. (2016). Sensitivity analysis of environmental models: A systematic review with practical workflow. *Environmental Modelling & Software*, *79*, 214–232. https://doi.org/10.1016/J.ENVSOFT.2016.02.008

Pianosi, F., Sarrazin, F., & Wagener, T. (2015). A Matlab toolbox for global sensitivity analysis. *Environmental Modelling & Software*, *70*, 80–85. https://doi.org/10.1016/J.ENVSOFT.2015.04.009

Pianosi, F., & Wagener, T. (2015). A simple and efficient method for global sensitivity analysis based on cumulative distribution functions. *Environmental Modelling & Software*, *67*, 1–11. https://doi.org/10.1016/J.ENVSOFT.2015.01.004

Potter, W. D. (1958). Upper and lower frequency curves for peak rates of runoff. *Transactions, American Geophysical Union*, *39*(1), 100. https://doi.org/10.1029/TR039i001p00100

Rogger, M., Pirkl, H., Viglione, A., Komma, J., Kohl, B., Kirnbauer, R., … Blöschl, G. (2012). Step changes in the flood frequency curve: Process controls. *Water Resources Research*, *48*(5). https://doi.org/10.1029/2011WR011187

Rohde, R., Muller, R., Jacobsen, R., Perlmutter, S., Rosenfeld, J., Wurtele, J., … Mosher, S. (2013). Berkeley earth temperature averaging process. *Geoinformatics & Geostatistics: An Overview*, *1*(2), 20–100. Retrieved from http://static.berkeleyearth.org/papers/Methods-GIGS-1-103.pdf

Sharma, A., Wasko, C., & Lettenmaier, D. P. (2018). If precipitation extremes are increasing, why aren't floods? *Water Resources Research* https://doi.org/10.1029/2018WR023749, *54*, 8545–8551.

Sikorska, A. E., Viviroli, D., & Seibert, J. (2015). Flood-type classification in mountainous catchments using crisp and fuzzy decision trees. *Water Resources Research*, *51*(10), 7959–7976. https://doi.org/10.1002/2015WR017326

Slater, L. J., & Villarini, G. (2016). Recent trends in U.S. flood risk. *Geophysical Research Letters*, *43*(24), 12,428–12,436. https://doi.org/10.1002/2016GL071199

Slater, L. J., & Villarini, G. (2017). On the impact of gaps on trend detection in extreme streamflow time series. *International Journal of Climatology*, *37*(10), 3976–3983. https://doi.org/10.1002/joc.4954

Smith, J. A., Cox, A. A., Baeck, M. L., Yang, L., & Bates, P. (2018). Strange floods: The upper tail of flood peaks in the United States. *Water Resources Research*, *54*(9), 6510–6542. https://doi.org/10.1029/2018WR022539

Spear, R. C., & Hornberger, G. M. (1980). Eutrophication in peel inlet—II. Identification of critical uncertainties via generalized sensitivity analysis. *Water Research*, *14*(1), 43–49. https://doi.org/10.1016/0043-1354(80)90040-8

Tarasova, L., Merz, R., Kiss, A., Basso, S., Blöschl, G., Merz, B., … Wietzke, L. (2019). Causative classification of river flood events. *Wiley Interdisciplinary Reviews: Water*, e1353. https://doi.org/10.1002/wat2.1353

Tockner, K., Uehlinger, U., & Robinson, C. (2009). *Rivers of Europe*. Retrieved from https://books.google.ch/books?hl=en&lr=&id=GDmX5XKkQCcC&oi=fnd&pg=PA1&dq=rivers+of+europe&ots=CSUhE8vZWU&sig=Q_xJY68LBITWLb1MC6mHLla0uYc

Viglione, A., Chirico, G. B., Komma, J., Woods, R. A., Borga, M., & Blöschl, G. (2010). Quantifying space-time dynamics of flood event types. *Journal of Hydrology*, *394*(1), 213–229. Retrieved from http://www.sciencedirect.com/science/article/pii/S0022169410003264

Villarini, G., & Slater, L. J. (2017). *Climatology of flooding in the United States* (ed., Vol. 1). https://doi.org/10.1093/acrefore/9780199389407.013.123

Vormoor, K., Lawrence, D., Heistermann, M., & Bronstert, A. (2015). Climate change impacts on the seasonality and generation processes of floods—Projections and uncertainties for catchments with mixed snowmelt/rainfall regimes. *Hydrology and Earth System Sciences*, *19*(2), 913–931. https://doi.org/10.5194/hess-19-913-2015

Wagener, T., & Wheater, H. S. (2006). Parameter estimation and regionalization for continuous rainfall-runoff models including uncertainty. *Journal of Hydrology*, *320*(1), 132–154. Retrieved from http://www.sciencedirect.com/science/article/pii/S0022169405003410

Wasko, C., & Sharma, A. (2017). Global assessment of flood and storm extremes with increased temperatures. *Scientific Reports*, *7*(1), 7945. https://doi.org/10.1038/s41598-017-08481-1

Waylen, P., & Woo, M. (1982). Prediction of annual floods generated by mixed processes. *Water Resources Research*, *18*(4), 1283–1286. https://doi.org/10.1029/WR018i004p01283

Whitfield, P. H. (2012). Floods in future climates: A review. *Journal of Flood Risk Management*, *5*(4), 336–365. https://doi.org/10.1111/j.1753-318X.2012.01150.x

Witten, I. H., Frank, E., & Hall, M. A. (2016). *Data mining: Practical machine learning tools and techniques*. https://doi.org/10.1016/c2009-0-19715-5

Young, P. C., Spear, R. C., & Hornberger, G. M. (1978). Modeling badly defined systems: some further thoughts. In *Proceedings of the SIMSIG Conference* (pp. 24–32). Canberra.

## SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

**How to cite this article:** Stein L, Pianosi F, Woods R. Event-based classification for global study of river flood generating processes. *Hydrological Processes*. 2019;1–16. https://doi.org/10.1002/hyp.13678